DTIC FILE COPY

AD-A222 409

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

2

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE 1982

3. REPORT TYPE AND DATES COVERED Reprint

4. TITLE AND SUBTITLE
ROBUST APPROXIMATIONS FOR THE FILTERING PROBLEM

5. FUNDING NUMBERS
AFOSR-86-0332
61102F 2304/A1

6. AUTHOR(S)
Robert J. Elliott

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Department of Statistics and Applied Probability
University of Alberta
Edmonton, Alberta, Canada
T6G 2G1

8. PERFORMING ORGANIZATION REPORT NUMBER
AFOSR-TR- 90-0589

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR/NM
Building 410
Bolling AFB, DC 20332-6448

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFOSR-86-0332

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release;
distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

A diffusion observation process is approximated by a Markov chain. The information obtained by observing the Markov chain is the same as that obtained by observing a related multivariate point process. Filtering and Zakai equations are obtained for multivariate point process observations. These involve Stieltjes integrals rather than Ito integrals with respect to Brownian motion, and so they provide robust formulae, that is, formulae which are continuous in the observation process. (KR)

14. SUBJECT TERMS

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAR

ROBUST APPROXIMATIONS FOR THE FILTERING PROBLEM

ROBERT J. ELLIOTT

Department of Statistics and Applied Probability
The University of Alberta
Edmonton, Alberta, Canada. T6G 2G1

ABSTRACT

A diffusion observation process is approximated by a Markov chain. The information obtained by observing the Markov chain is the same as that obtained by observing a related multivariate point process. Filtering and Zakai equations are obtained for multivariate point process observations. These involve Stieltjes integrals rather than Ito integrals with respect to Brownian motion, and so they provide robust formulae, that is, formulae which are continuous in the observation process.

1. FILTERING

All processes will be defined on a complete probability space (Ω, F, P). The time parameter $t \in [o,\infty)$ and there is a right continuous, complete filtration $\{F_t\}$ on Ω with respect to which all processes are adapted. Consider two independent Brownian motions $\{B_t\} = \{B_t^1, \ldots, B_t^M\}$, $\{W_t\} = \{W_t^1, \ldots, W_t^n\}$, and suppose the signal process is the solution of the stochastic differential equation:

$$dx_t = f(x_t)dt + \sigma(x_t)dB_t$$
$$x_o \in R^d. \qquad (1)$$

Acknowledgements: This work was supported by the U.S. Army Research Office under Contract DAAL03-87-K-0102 and by the Night Vision and Electro-Optics Center, and Battelle, Research Triangle Park, N.C. under contract DAAG29-81-D-0100.

The signal is not observed directly but by means of the observation process $\{y_t\}$, where $y_t \in R^n$ is the solution of:

$$dy_t = h(x_t)dt + dw_t$$
$$y_o = 0 \in R^n. \qquad (2)$$

For simplicity suppose that f, σ and h are bounded and continuous, and the solution of (1) is unique in the sense of probability distributions. The following formulation of the non-linear filtering problem is developed in detail in [1]. Write $Y_t^o = \sigma\{y_s : s \leq t\}$ for the σ-field representing the history of the observation process up to time t and $Y = \{Y_t\}$ for the right continuous complete filtration generated by Y_t^o. The best estimate in, say, the mean square sense of x_t, given the history Y_t is $E[x_t|Y_t]$. To determine this conditional distribution it is enough to find $E[\emptyset(x_t)|Y_t]$ for any twice, (or even infinitely), differentiable real valued function $\emptyset$ with compact support on R^d. Using the differentiation rule

$$\emptyset(x_t) = \emptyset(x_o) + \int_o^t L\,\emptyset(x_s)ds + \int_o^t \nabla\emptyset\cdot\sigma(x_s)dB_s$$

where L is the second order operator

$$L = \sum_{i,j=1}^m a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^m f_i(x)\frac{\partial}{\partial x_i}$$

with $a = (a_{ij}) = \sigma\sigma'/2$.

Write π for the Y-optional projection so that $\pi_t(\emptyset) = \pi(\emptyset(x_t)) = E[\emptyset(x_t)|Y_t]$. Then the filtering equation [1] for $\pi_t(\emptyset)$ is:

$$\pi_t(\emptyset) = \pi_o(\emptyset) + \int_o^t \pi_u(L\emptyset)du + \int_o^t (\pi_u(\emptyset h) - \pi_u(\emptyset)\,\pi_u(h))'d\nu_u \qquad (3)$$

Here ν, the innovations process, is a (Y,P) Brownian motion defined by

$$\nu_t = y_t - \int_0^t \pi(h)_u \, du \qquad (4)$$

Difficulties with Equation (3) are that it is quadratic in π and involves an Ito stochastic integral with respect to ν.

2. APPROXIMATIONS FOR THE OBSERVATIONS

In this section we shall describe an approximation to the observation process y by a Markov chain. Our approximation to the observation process is similar to that for diffusions introduced by Kushner [2]. The resulting (approximate) filtering equations involve Stieltjes integrals, rather than Ito integrals as in (3). Why should approximations to the observation process be considered? This question can be countered by pointing out the idealized nature of a diffusion, and by noting that all measurements are approximations, so that at some level of accuracy any observation process is a jump process. Alternatively we could just say we wish to consider observations of this form.

To approximate the process $\{y_t\}$ by a Markov chain y^δ consider $\delta>0$ and the grid R^n_δ on R^n with difference parameter δ. Write

$$Q_\delta(x_t) = n + \delta \sum_{i=1}^n |h_i(x_t)|,$$

$$\Delta t^\delta(x_t) = \delta^2/Q_\delta(x_t).$$

Write ϵ_i for the unit vector in the i^{th} coordinate direction of R^n, and for $y \in R^n_\delta$, $y' = y \pm \delta \epsilon_i$ define:

$$P^\delta(y, y', x_t) = (1/2 + \delta h_i^{\pm}(x_t))/Q_\delta(x_t). \qquad (5)$$

Then $P^\delta(y, y', x_t)$ will be the probability of a jump from y to y', given x_t and given there is a jump at time t. Also, $P^\delta(y, y', x_t) \geq 0$ and $\sum_{y'} P^\delta(y, y', x_t) = 1$.

The conditional distributions of the jump times of the approximating Markov process $\{y_t^\delta\}$ are defined by:

$$P\,(\text{next jump after } t + s \mid \text{previous jump at } t \text{ and } x_u,\ t \leq u \leq t + s) =$$

$$\exp\left(- \int_s^{s+t} \frac{du}{\Delta t^\delta(x_u)}\right). \qquad (6)$$

The signal $\{x_t\}$ enters the Markov chain $\{y^\delta\}$ through (5) and (6). We can then establish

LEMMA 2.1. If $y' = y \pm \delta \epsilon$ is the next state value $E[y - y' | x_t] = h(x_t)\, \Delta t^\delta(x_t)$ and $\mathrm{Cov}[y' - y | x_t] = I \Delta t^\delta(x_t) + o(\Delta t^\delta(x_t))$ where I is $n \times n$ identity matrix.

REMARKS 2.2. The approximating Markov chain $\{y^\delta\}$, with state space R^n_δ, is determined by a sequence of jump times $\{T_i\}$, $i = 1, 2, \ldots\ldots$ and a sequence of positive or negative unit vectors $\{\pm \epsilon_i\}$ in R^n which describes the jump values. That is, the Markov chain can be thought of as a marked point process $\{(T_n, Z_n)\}$, with jump times T_n and jumps Z_n in the state space $\{\pm \epsilon_i\}$. In turn, this marked point process can be considered as a multivariate point process

$$N_t = \{N_t(+1), N_t(-1), \ldots, N_t(-n)\}$$

where $N_t(\pm i) = \sum_{n \geq 1} I(t \geq T_n)\, I(Z_n = \pm \epsilon_i)$

The distribution of the jump times T_n is given by (6), so roughly

$$P(T_{n+1} = t + s | T_n = t,\ T_{n+1} \geq t + s \text{ and } x_u,\ t \leq u \leq t + s)$$

$$= (\Delta t^\delta(x_{t+s}))^{-1} = Q_\delta(x_{t+s})/\delta^2.$$

Also, $P(Z_n = \pm \epsilon_i | T_n = t \text{ and } x_t)$

$= (1/2 + \delta h_i^{\pm}(x_t))/Q_\delta(t)$. These two conditional probabilities determine the Levy system (see [1]) of the approximating Markov chain y^δ. Write

$$\lambda_{\pm i}(t) = (1 + 2 \delta h_i^{\pm}(x_t))/2\delta^2.$$

Then we can show the processes

$$Q_t^{\pm i} = N_t(\pm i) - \int_0^t \lambda_{\pm i}(u)\, du$$

are $\{F_t\}$ martingales. The $\lambda_{\pm i}$ are, therefore, the intensities (see [1] of the point processes $N(\pm i)$.

3. FILTERING WITH MULTIVARIATE POINT PROCESS OBSERVATIONS

We have described an approximation to the original filtering problem in which the signal modulates a

multivariate point process

$N_t = \{N_t(\pm i)\ ,\ i = 1, \ldots, n\}$. Write

$$Y_t^{o\delta} = \sigma\{N_s : s \leq t\}$$

and $Y^\delta = \{Y_t^\delta\}$ for the right continuous, complete filtration generated by the $Y_t^{o\delta}$. Suppose $\{\phi_t\}$ is a process. Then $\{\hat{\phi}_t\}$ will denote the Y^δ optional projection of $\{\phi_t\}$ and $\{\overset{*}{\phi}_t\}$ will denote the Y^δ predictable projection. Then we have the fundamental result:

LEMMA 3.1. $\hat{Q}_t^i = N_t(i) - \int_0^t \overset{*}{\lambda}_i(u)du$ is a Y^δ-martingale for $i = \pm 1, \ldots, \pm n$.

The filtering equation for $\phi(x_t)$ with observation process N is proved similarly to that in [1], and states the following:

THEOREM 3.2 $\hat{\phi}(x_t) = \hat{\phi}(x_o)$

$$+ \int_0^t \widehat{L\phi}\, du + \sum_{i=\pm 1}^{\pm n} \int_0^t \gamma_u^i\, d\hat{Q}_u^i \qquad (7)$$

where

$$\gamma_u^i = (\widehat{\phi(x_u)\ \lambda_i(u)} - \hat{\phi}(x_u)\ \overset{*}{\lambda}_i(u))(\overset{*}{\lambda}_i(u))^{-1}$$

REMARKS 3.3. The Ito stochastic integral in (3) is now replaced by the (robust) Stieltjes integrals with respect to the $\hat{Q}^i$. However, the γ^i still involve a product of projections and/or a division by $\overset{*}{\lambda}_i(u)$. These difficulties can be circumvented by considering a different probability measure P_1.

Suppose that under P_1, (the reference probability), each component $N(i)$ of the multivariate point process

$$N = (N(+1), (N(-1), \ldots, N(+n), N(-n))$$

is a standard Poisson process of intensity 1. Then for $i = \pm 1, \ldots, \pm n$

$$\bar{Q}_t^i = N_t(i) - t \qquad (8)$$

is a Y^δ- martingale under P_1. Consider the family of exponentials

$$\Lambda_t = 1 + \sum_{i=\pm 1}^{\pm n} \int_0^t \Lambda_{u-}(\lambda_i(u)-1)\ d\bar{Q}_u^i. \qquad (9)$$

Then Λ is an (F, P_1) martingale, $\Lambda_t > 0$ a.s. and $E[\Lambda_t] = 1$. Define a new probability measure $P = P_\lambda$ by

$$E_1[\frac{dP}{dP_1}|F_t] = \Lambda_t.$$

Here E_1 denotes expectation with respect to P_1. An application of Baye's rule establishes the following result:

LEMMA 3.4. Under $P = P_\lambda$ the processes

$$Q_t^i = N_t(i) - \int_0^t \lambda_i(u)\ du,$$

$i = \pm 1, \ldots, \pm n$ are $\{F_t\}$- martingales.

NOTATION 3.5. Write $\hat{\Lambda}$ for the Y^δ-optional projection of Λ under P_1. Then for each $t \geq 0$ $\hat{\Lambda}_t = E_1[\Lambda_t | Y_t^\delta]$a.s. and $\hat{\Lambda}$ is a locally square integrable (Y^δ, P_1) martingale. Because we now have the two measures P_1 and $P = P_\lambda$, for any process $\zeta = \{\zeta_t\}$ we shall write $\pi(\zeta)$, resp. $\pi_p(\zeta)$, for the (Y^δ, P) optional, resp. predictable, projection of ζ. Therefore,

$$\pi_p(\lambda_i(u)) = E[\lambda_i(u)|Y_{u-}^\delta] \text{ a.s.}$$

for $i = \pm 1, \ldots, \pm n$.

LEMMA 3.6. $$\hat{\Lambda}_t = 1 + \sum_{i=\pm 1}^{\pm n} \int_0^t \phi_u^i\ d\bar{Q}_u^i \qquad (10)$$

where $\phi_u^i = \hat{\Lambda}_{u-}(\pi_p(\lambda_i(u)) - 1)$.

REMARKS 3.7. Using the π notation, Theorem 3.2 states:

$$\pi(\phi(x_t)) = \pi(\phi(x_o)) + \int_0^t \pi(L\phi(x_u))du + \sum_{i=\pm 1}^{\pm n} \int_0^t \gamma_u^i\ d\hat{Q}_u^i \qquad (11)$$

where

$$\gamma_u^i = (\pi_p(\phi(x_u)\ \lambda_i(u)) - \pi(\phi(x_u))\ \pi_p(\lambda_i(u))\ \pi_p(\lambda_i(u))^{-1}.$$

Now by Baye's formula:

$$\pi(\emptyset(x_t)) = E[\emptyset(x_t)|Y_t^\delta]$$
$$= E_1[\Lambda_t\emptyset(x_t)|Y_t^\delta]\,(E_1[\Lambda_t|Y_t^\delta])^{-1}$$
$$= \sigma(\emptyset(x_t))\,\sigma(1)^{-1},$$

where $\sigma(\emptyset(x_t)) = E_1[\Lambda_t\ \emptyset(x_t)|Y_t^\delta]$

is an unnormalized conditional expectation and $\sigma(1) = \hat{\Lambda}_t$.

Therefore, $\sigma(\emptyset(x_t)) = \hat{\Lambda}_t\ \pi(\emptyset(x_t))$ and by computing the product of (10) and (11) we obtain the following Zakai equation for $\sigma(\emptyset(x_t))$:

THEOREM 3.8. $\sigma(\emptyset(x_t)) = \sigma(\emptyset(x_o)) +$

$$\int_o^t \sigma(L\emptyset(x_u))\,du + \sum_{i=\pm 1}^{\pm n}\int_o^t \sigma(\emptyset(x_u)(\lambda_i(u) - 1))\,d\bar{Q}_u^i. \tag{12}$$

REMARKS 3.9. The advantages of equation (12) are that it is linear in σ and, because it involves only Stieltjes integrals with respect to the Poisson processes $\bar{Q}^i$, it is robust.

The remaining questions concern the convergence of the approximate filters as the mesh size δ goes to zero. In fact it can be shown that the family of filtered processes given by Theorem 3.2. for $\delta>0$ is tight and converges weakly to the process given by (3) as $\delta\to 0$.

However, the family of processes given by the Zakai equation (12) does not converge as $\delta\to 0$. To obtain a convergent set of processes one must assume that for each $\delta>0$ there is a multivariate point process $N_t^\delta = (N_t^\delta(+1), N_t^\delta(-1), \ldots, N_t^\delta(-n))$ such that under the reference measure P_1 each component is a point process with rate t/δ^2. Λ_t^δ is defined by:

$$\Lambda_t^\delta = 1 + \sum_{i=\pm 1}^{\pm n}\int_o^t \Lambda_u^\delta\,(2\delta^2\,\lambda_i(u) - 1)\,d\bar{Q}_u^{i\delta} \tag{13}$$

where $\bar{Q}_t^{i\delta} = N_t^\delta(i) - t/\delta^2$.

A new probability measure P^δ is given by:

$$E_1[\frac{dP^\delta}{dP_1}\,|F_t] = \Lambda_t^\delta.$$

The Zakai equation for:

$$\sigma(\emptyset(x_t)) = E_1[\emptyset(x_t)\ \Lambda_t^\delta|Y_t^\delta]$$

is then:

$$\sigma(\emptyset(x_t)) = \sigma(\emptyset(x_o)) + \int_o^t \sigma(L\emptyset(x_u))\,du + \bar{2}\sum_{i=\pm 1}^{\pm n}\int_o^t \sigma(\emptyset(x_u)\ h_i)\delta\ d\bar{Q}_u^{i\delta}. \tag{14}$$

Now the family of processes given by (14) for $\delta>0$ is tight and converges weakly as $\delta\to 0$ to a process $\sigma(\emptyset(x_t))$ which is the solution of:

$$\sigma(\emptyset(x_t)) = \sigma(\emptyset(x_o)) + \int_o^t \sigma(L\emptyset(x_u))\,du + 2\sum_{i=\pm 1}^{\pm n}\int_o^t \sigma(\emptyset(x_u)\ h_i(x_u))dw_u^i$$

where $w = (w^{+1}, w^{-1}, \ldots, w^{+n}, w^{-n})$ is a Brownian motion.

REFERENCES

1. R.J. Elliott, "Stochastic calculus and applications", Springer-Verlag, New York-Hiedelberg-Berlin, 1982.

2. H.J. Kushner, "Probability methods for approximations in stochastic control and for elliptic equations", Academic Press, New York, 1977.